Report #5
PD-Internship
Lerner Paul
25/04/2018

Before April 8th I hadn't extensively searched for papers about RNNs. So I did with the following search equations on Google Scholar :

- "Long Short Term Memory"
- "LSTM"
- "Gated Recurrent Unit"
- "GRU"
- "Recurrent Neural Network"
- "RNN"

& :

- "Handwriting" & ¬"Recognition"
- "Online Handwriting" (also by itself so I would not only have works about RNNs)
- "Drawing"
- "Dysgraphia"
- "Tablet"
- "Micrographia"
- <Disease Name> (cf. Report #3)

Thanks to this search I was able to find new works on Neurodegenerative Diseases (NDs) so I updated Report #3 (I added annotations to the pdf on the updates).

Summary

1 Online Handwriting Analysis

1.1 Literature review

Handwriting analysis is referred to as online when the studied data is recorded through a tablet or smart pen, as opposed to offline where we study the strokes of pen on paper.

Cf. Report #2 and Report #3 for CDSS based on HaW analysis.

Likforman-Sulem et al. compared **Random Forest** classifier as well as an **SVM** to assess one's emotional state (i.e. depression, anxiety and stress).

Mekyska et al. use a **Random Forest** classifier to assess children's dysgraphia. They compare it with a **Linear Discriminant Analysis** and obtain better results with random Forest. The workflow and extracted features are similar to those of Drotar et al. (more or less the same group of authors). Rosenblum & Dror use however quite different features and a **SVM** for the same task. They achieve 90 % Se and Sp though on a different dataset of Mekyska et al. who achieve 96% Se and Sp.

Graves et al. used a **bLSTM** for online handwriting recognition. Their model is quite similar to ours : They standardize the data, have hidden layers of 100 units. They train their model with a learning rate of 10^{-4} for a maximum of 50 epochs (if no early stopping). They outperform SoA's HMM. They use IAM-onDB database.

Zhang et al. 2017 use a **bLSTM** for writer identification. They segment their data into subsequences of 100 timesteps. The prediction for all subsequences are made independently then combined (late fusion) for the final writer ID. Instead of using majority voting to merge the models prediction they average over the outputs (i.e. the probabilities of each class). As their original data sequences are very long (6k and 9k in average for both their dataset, respectively) they randomly sample 1000 subsequences from each original sequence. The sampled subsequences may overlap. Even if they don't, this approach may allow for *attention* layers as the input size is fixed (100 timesteps here). They leave it for future work. It's interesting to note that they used random sampling because of the context of their task (writer id may be required at any given time of the sequence) but obtained similar results with fixed-position sampling.
They outperform SoA with 100% and 99% accuracy on the English dataset (133 subjects) and Chinese dataset (186 subjects), respectively.
They transform each sequence S from the data into ΔS as in the equations below so that it represents the direction of the pen movement.

$$S = [[x_1, y_1, p_1], [x_2, y_2, p_2], \ldots, [x_n, y_n, p_n]]$$

$$\begin{aligned}\Delta S = [[&x_2 - x_1, y_2 - y_1, p_2 \times p_1], \ldots \\ &[x_i - x_{i-1}, y_i - y_{i-1}, p_i \times p_{i-1}], \ldots \\ &[x_n - x_{n-1}, y_n - y_{n-1}, p_n \times p_{n-1}]]\end{aligned}$$

With p_* being the button status.
They sum the 2 hidden states from both directions before feeding it to a "logistic regression layer". I find that very strange, I wonder if they meant a FC layer because they say they train their model using multiclass negative log-likelihood loss (i.e. cross-entropy).
They also init the forget gate bias at large value (here 5) as advocated by Jozefowicz et al. (cf. Report #4).

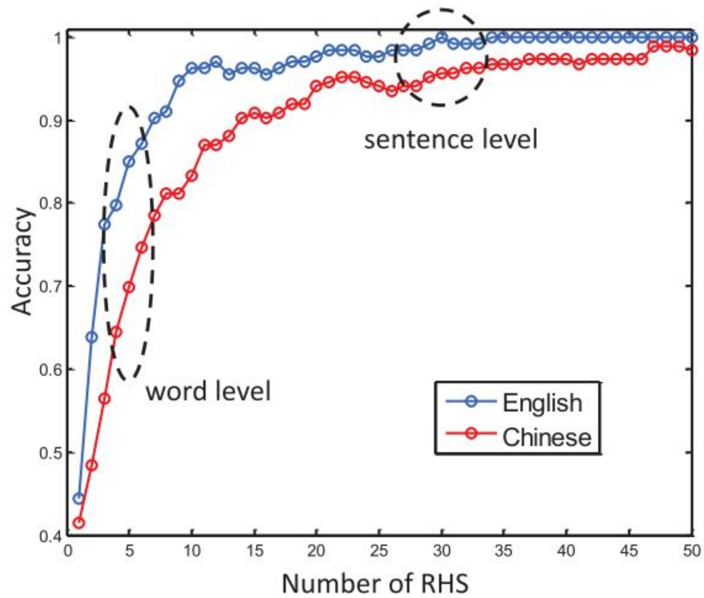

Fig. from Zhang et al. 2017. Accuracy VS number of subsequences (RHS) used to make final prediction.
We can see that the accuracy goes up very fast until ~10 subsequences and reaches maximum for ~30 and ~50 subsequences for english and chinese, respectively.
They are able to perform Cross Learning between the english and chinese dataset, yielding 33% accuracy when going from english to chinese and 81% when going from chinese to english. Moreover, when training on the both set at the same time they achieve 96% accuracy.

1.2 Motivation for using RNN - Conclusion

RNNs such as LSTMs and GRUs have been used successfully for Handwriting recognition both online and offline (Graves et al.), mode detection (Indermühle et al. 2012), writer identification (Zhang et al. 2017) and number of other time-series classification task (cf. Report #4) ; although they're not SoA for Online HaW recognition (Keysers et al.). For PD diagnosis, RNNs have been used based on EEG and EMG signals (cf. Report #3). They haven't been for PD diagnosis based on HaW (cf. Report #2) nor for any other ND diagnosis based on HaW (cf. Report #3).
RNNs can take as input raw data as well as hand-crafted features. E.g. Graves et al. use raw data for online handwriting recognition and Indermühle et al. 2012 extract features for mode detection. Both use a bLSTM.
One of major shortcomings of hand-crafted features is that it requires expert knowledge and are often specific to the data. See, e.g. Drotar et al. in Report #2 who had quite poor results on the spiral task because they used text-specific features. Unlike Moetesum et al.
Zhang et al. empirically demonstrated the generalization power of RNNs by cross learning between a Chinese and an English handwriting dataset for writer identification (cf. 1.1 Literature review).

2 Online Handwriting and Drawing Datasets

We could use these datasets for transfer learning purposes :

Authors	**Name**	**Size**	**Sampling Rate**	**Content**	**Measures**
Guyon et al.	UNIPEN	5 million characters / 490 MB	100Hz	English script	x and y coordinates
Liwicki M. and Bunke.	IAM-OnDB	86 272 words / 440 MB	?	English script	x, y and timestamp
Indermühle et al. 2010	IAMonDo-database	68 841 words + 1000s of	?	English script and drawings	x, y, timestamp and "force"

		drawings / 126 MB			(pressure ?)

Other datasets :

- Alonso-Martinez et al. mention different (small) online handwriting datasets.
- Likforman-Sulem et al. — EMOTHAW.
- Liu et al. — CASIA Online and Offline Chinese Handwriting Databases.

References

Alonso-Martinez, C., Faundez-Zanuy, M., & Mekyska, J. (2017). A Comparative Study of In-Air Trajectories at Short and Long Distances in Online Handwriting. Cognitive Computation, 9(5), 712.

Graves, A., Liwicki, M., Bunke, H., Schmidhuber, J., & Fernández, S. (2008). Unconstrained on-line handwriting recognition with recurrent neural networks. In Advances in neural information processing systems (pp. 577-584).

Guyon, I., Schomaker, L., Plamondon, R., Liberman, M., & Janet, S. (1994, October). UNIPEN project of on-line data exchange and recognizer benchmarks. In *Proceedings of the 12th IAPR International Conference on Pattern Recognition, Vol. 3-Conference C: Signal Processing (Cat. No. 94CH3440-5)* (Vol. 2, pp. 29-33). IEEE.

Indermühle, E., Liwicki, M., & Bunke, H. (2010, June). IAMonDo-database: an online handwritten document database with non-uniform contents. In *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems* (pp. 97-104). ACM.

Indermühle, E., Frinken, V., & Bunke, H. (2012, September). Mode detection in online handwritten documents using BLSTM neural networks. In 2012 International Conference on Frontiers in Handwriting Recognition (pp. 302-307). IEEE.

Keysers, D., Deselaers, T., Rowley, H. A., Wang, L. L., & Carbune, V. (2017). Multi-language online handwriting recognition. IEEE transactions on pattern analysis and machine intelligence, 39(6), 1180-1194.

Likforman-Sulem, L., Esposito, A., Faundez-Zanuy, M., Clémençon, S., & Cordasco, G. (2017). EMOTHAW: A novel database for emotional state recognition from handwriting and drawing. *IEEE Transactions on Human-Machine Systems*, *47*(2), 273-284.

Liu, C. L., Yin, F., Wang, D. H., & Wang, Q. F. (2011, September). CASIA online and offline Chinese handwriting databases. In *2011 International Conference on Document Analysis and Recognition* (pp. 37-41). IEEE.

Liwicki M. and Bunke H. IAM-OnDB - an on-line English sentence database acquired from handwritten text on a whiteboard. In Proc. 8th Int. Conf. on Document Analysis and Recognition, volume 2, pages 956–961, 2005.

Mekyska, J., Faundez-Zanuy, M., Mzourek, Z., Galaz, Z., Smekal, Z., & Rosenblum, S. (2017). Identification and rating of developmental dysgraphia by handwriting analysis. IEEE Transactions on Human-Machine Systems, 47(2), 235-248.

Rosenblum, S., & Dror, G. (2017). Identifying developmental dysgraphia characteristics utilizing handwriting classification methods. IEEE Transactions on Human-Machine Systems, 47(2), 293-298.

Sesa-Nogueras, E., Faundez-Zanuy, M., & Mekyska, J. (2012). An information analysis of in-air and on-surface trajectories in online handwriting. *Cognitive Computation*, *4*(2), 195-205.

Zhang, X. Y., Xie, G. S., Liu, C. L., & Bengio, Y. (2017). End-to-end online writer identification with recurrent neural network. IEEE Transactions on Human-Machine Systems, 47(2), 285-292.